Computers in Biology and Medicine 43 (2013) 1955–1964

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

The impact of noise on the reliability of heart-rate variability and complexity analysis in trauma patients

Nehemiah T. Liu*, Andriy I. Batchinsky, Leopoldo C. Cancio, José Salinas

U.S. Army Institute of Surgical Research, Fort Sam Houston, TX, United States

ARTICLE INFO

Article history:
Received 7 May 2013
Accepted 16 September 2013

Keywords:
Signal detection analysis
Heart rate variability
Heart rate complexity
Clinical decision support systems
Automatic data processing
Electrocardiogram

ABSTRACT

This study focused on the impact of noise on the reliability of heart-rate variability and complexity (HRV, HRC) to discriminate between different trauma patients and to monitor individual patients. Life-saving interventions (LSIs) were chosen as an endpoint because performance of LSIs is a critical aspect of trauma patient care. Noise was modeled and simulated by modifying original R–R interval (RRI) sequences via decimation, concatenation, and division of RRIs, as well as R-wave detection using the electrocardiogram. Results showed that under increasing simulated noise, entropy and autocorrelation measures can still effectively discriminate between LSI and non-LSI patients and monitor individuals over time.

1. Introduction

Despite extensive studies in the past few decades [1–4] on the significance and meaning of the many different measures of heart-rate variability (HRV) and heart-rate complexity (HRC), few studies exist to assess the influence of noise on these measures, and even less, on measures for triage and treatment of trauma patients. There is no doubt that noise plays an integral part in the clinical environment, but to what extent it affects the measures of HRV and HRC so as to misconstrue their meaning for clinical use and patient diagnosis remains a profound and unanswered question.

HRV and HRC metrics are used to quantify beat-to-beat changes in the R-to-R interval (RRI) of the electrocardiogram (ECG), and therefore, reflect different physiological factors modulating normal sinus rhythm [1–4]. Because the calculation of ECG-derived metrics relies on accurate R-wave detection (RWD) as well as ECG waveform data acquisition, a hurdle in using HRV and HRC as new vital signs is that various noise and artifact types must be accounted for at every stage prior to metric calculations. If HRV and HRC indices become unreliable or uninformative due to noise at any stage in the process, then use of these indices must be avoided during patient assessment.

To define noise in a medical setting, however, is complex (and sometimes, controversial), since ECGs, like other commonly monitored physiological signals, are often corrupted by artifacts, missing data, and noise that is non-Gaussian, non-linear, and non-stationarity [5]. In addition to electromechanical noise (motion artifacts), most ECGs contain atypical phenomena, i.e., arrhythmic changes consistent with illness, severe exsanguinations, and pre-terminal and terminal loss of ECG morphology. Furthermore, RWD can be impeded by baseline wander within the ECG as well as high frequency and electromyogram noise and/or power-line interference [6–8]; other possible impediments include human interaction with patients, medical devices, ECG leads, etc.

Because many potential sources can contribute to poor ECG signal quality and RWD performance, noise was defined for this study at the metric level rather than at the waveform and detection levels. In other words, noise is any source that alters the true RRI sequence of an ECG, and consequently, modifies calculation of an ECG-derived metric, the “gold standard” being determined by manual verification of the RRI sequence and subsequent metric calculations. This broad definition includes missed beats, false detections, and missing data, and thus, captures inherent weaknesses in existing RWD algorithms. If the altered RRI sequence is considered to be the result of a manually verified sequence having passed through some noisy channel, then an additive noise model from signal processing theory may best describe the above definition. This channel may be a cascade of filters, such as decimation and threshold filters that preclude specific intervals followed by interpolation filters that insert extra RRIs into the original RRI sequence.

In view of recent work by Proctor et al. [9] reviewing clinical application of HRV and HRC for triage and assessment of trauma

* Correspondence to: U.S. Army Institute of Surgical Research, 3650 Chambers Pass, Building 3610, Fort Sam Houston, TX 78234-6315, United States. Tel.: +210 539 2459; fax: +210 539 6244.
E-mail addresses: nehemiah.liu@us.army.mil (N.T. Liu), andriy.batchinsky1@us.army.mil (A.I. Batchinsky), lee.cancio@us.army.mil (L.C. Cancio), jose.salinas4@us.army.mil (J. Salinas).

0010-4825/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.compbiomed.2013.09.012

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
01 NOV 2013	**N/A**	**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
The impact of noise on the reliability of heart-rate variability and complexity analysis in trauma patients	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
Liu N. T., Batchinsky A. I., Cancio L. C., Salinas J.,	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
United States Army Institute of Surgical Research,JBSA Fort Sam Houston, TX	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**UU**	**10**	

patients, as well as Moorman et al. [10,11] and Seely et al. [12,13] investigating the clinical use of HRV and HRC for detecting sepsis and multiorgan failure, reliability of metric calculations may help detect significant changes from baseline values earlier and more accurately. Moreover, improved HRV and HRC calculations could help improve trends over seconds, minutes, or even days and identify crossed thresholds that would have otherwise been missed due to noise and/or poor RWD performance. Consequently, improved HRV and HRC accuracies could enhance clinical decision making as well as decision support systems and patient care in the areas mentioned above. Other applications which require monitoring over time, such as mentioned in [9,14], could likewise benefit from reliable HRV and HRC values.

This study focused on the impact of noise on the reliability of HRV/HRC analysis in trauma patients, with life-saving interventions (LSIs) as an endpoint, because effective and timely performance of LSIs is a critical aspect of trauma patient care. The purpose of this study was to investigate how simulated noise (i.e., artificial modifications of true RRI sequences) could potentially affect selected measures of HRV and HRC in their usefulness for discriminating between trauma patients who received at least one LSI and those who received none, and in their usefulness for monitoring trauma patients in real time. A hypothesis of this study was that the calculation of certain metrics may be less susceptible to noise, and therefore, may be integrated into a real-time software program for decision support and triage in critically ill and trauma patients.

2. Materials and methods

2.1. Patient data

This study was conducted under a protocol reviewed and approved by the U.S. Army Medical Research and Materiel Command Institutional Review Board, and in accordance with the approved protocol. Following approval, 108 pre-hospital patient records were selected from the U.S. Army Institute of Surgical Research Trauma Vitals (TV) database based upon the availability of ECG waveform data and manual verification of all RRI sequences. Because all data were analyzed post hoc, the study was considered minimal risk, and informed patient consent was waived.

Importantly, all ECGs were acquired at a sampling frequency of 375 Hz using a Welch Allyn PIC 50 (Welch Allyn, Skaneateles Falls, NY) monitor and characterized the underlying heart rhythms of severe trauma patients (Code 2/3) with blunt and penetrating injuries, who were transported from the scene by helicopter service to a Level I trauma center in Houston, TX or San Antonio, TX. Of these 108 waveforms, 82 ECGs (76%) belonged to patients who received at least one LSI, while the remaining 26 ECGs (24%) belonged to patients who received none. Waveform data were extracted by research personnel and uploaded to the TV database for analysis using a personal digital assistant attached to the monitor during transport. Lengths of patient ECGs varied from approximately 15 to 20 min.

2.2. R–R interval computation

The “gold standard” for obtaining true RRI sequences was manual verification of R waves, which was accomplished by importing ECG waveform data into WinCPRS software (Absolute Aliens Oy, Turku, Finland), visually analyzing the data, and marking times and positions of all R waves.

Because many previously published results of beat detection algorithms against different databases (e.g., MIT–BIH Arrhythmia Database) involved the detection of beats or QRS complexes, rather than the detection of R-waves, these results may not reflect stringent requirements on performance. In other words, beat detection algorithms may not be accurate enough for the real-time calculation of measures of HRV and HRC. Therefore, ECGs of patient records were loaded into an in-house-developed real-time RWD software algorithm [15] in order to produce detected RRI sequences. This software algorithm was a data fusion algorithm that employs four individual RWD algorithms to detect R waves in the ECG. Central to its fusion scheme, the data fusion algorithm selects the mode RRI or the RRI closest to a previous averaged decision within a given time frame. A strict tolerance of 25 ms was used for classifying detected R waves as true positives, since 25 ms corresponded to within 3% of an average RRI of 750 ms (or a heart rate of 80 beats per minute). Detected beats not satisfying this criterion were classified as false positives; missed beats were classified as false negatives.

2.3. Noise modeling

Increasing levels of noise may mask acute changes in a trauma patient's status. Because the methods for quantitatively measuring the effects of noise on the reliability of HRV and HRC metrics are still yet to be determined from a clinical standpoint, noise was modeled and simulated using four techniques: (1) decimation of an RRI sequence at selected multiples of the original sequence, (2) concatenation of adjacent RRIs at selected multiples of the original sequence, (3) division of RRIs into two RRIs at selected multiples of the original sequence and (4) detection of an RRI sequence using the ECG waveform. Selected multiples of the original sequence denote only those Mth multiples of the true RRI sequence, for $M=10, 9, \ldots, 2$. Hence, for every true original RRI sequence, an additional $9+9+9+1=28$ RRI sequences were produced. In the case of (3), division of RRIs by two was only performed when the resulting RRIs were each greater than 250 milliseconds in order to ensure that newly created beats did not fall within a refractory period of 220 milliseconds.

Decimation was chosen to partially randomize a true sequence in a controlled manner, thereby ensuring that newly formed sequences could be related to the original sequence. It was also used to simulate missing or dropped data. However, as this technique did not generate sequences for an exhaustive analysis, concatenation of selected RRIs was used to simulate missed beats and arrhythmias, while division of selected RRIs was used to simulate ectopic beats, motion artifacts, and arrhythmias. Therefore, increasing the value of M meant increasing the randomness, number of missed beats, or number of ectopic beats and motion artifacts in a sequence, respectively. $1/M \times 100$ equaled the percentage that a true RRI sequence was affected by simulated noise.

2.4. Analysis of heart-rate variability and complexity measures

For this study, nine metrics were selected based upon their clinical relevance, frequent citations in the biomedical research

Table 1
Selected measures of heart-rate variability and complexity.

Measure	Notation
Sample entropy	SampEn
Quadratic sample entropy	QSE
Multiscale entropy	MSE
Poincaré variability ratio	SD1/SD2
Fractal scaling exponent	α
Autocorrelation coefficient	$A(k, \tau)$
Degree of non-stationarity	StatAv
Standard deviation	SDRR
Successive differences	SDSD

Table 2
Demographics of selected patients from the trauma vitals database.

Variable	All patients		Patients with life-saving interventions (LSIs)		Number of LSIs
	Number (*N*)	Percentage (*N*/108)	Number (*n*)	Percentage (*n*/*N*)	
All patients	108	100	82	76	142
Gender					
Female	25	23	19	76	39
Male	82	76	62	76	102
Unknown	1	1	1	100	1
Race					
White/Caucasian	44	41	31	70	58
Black	6	6	5	83	9
Hispanic	24	22	21	88	34
Asian/Pacific	3	3	3	100	6
Not recorded	31	28	22	71	35
Age					
Mean 37 ± 14					
Quartiles					
18–24	25	23	21	84	35
25–35	29	27	23	79	38
36–45	27	25	17	63	27
46–83	27	25	21	78	42
Method of injury					
Blunt	93	86	73	78	126
Penetrating	13	12	9	69	16
Unknown	2	2	0	0	0
Injury severity					
Mean 17 ± 11					
Quartiles					
1–8	19	18	11	58	15
9–12	26	24	16	62	25
13–22	31	29	26	84	49
24–59	32	29	29	91	53
Systolic BP[a]					
Mean 116 ± 23					
Median 115.5					
80–89	7	6	6	86	11
90–119	29	28	24	83	41
120–129	8	7	6	75	7
130–138	8	7	3	38	7
140–170	12	11	9	75	12
Unknown	44	41	34	77	64
Diastolic BP[a]					
Mean 79 ± 16					
Median 80					
44–58	7	6	6	86	11
60–76	17	16	14	82	23
80–87	9	8	7	78	10
90–98	14	13	7	50	12
100–110	4	4	3	75	5
Unknown	57	53	45	79	81
Heart rate[a]					
Mean 106 ± 27					
Median 104					
20–56	3	3	1	33	1
76–98	24	22	18	75	23
100–108	10	9	7	70	17
110–116	7	6	5	71	7
120–127	9	8	8	89	16
131–156	12	11	9	75	15
Unknown	43	41	34	79	63

BP, blood pressure (mm Hg); heart rate (beats per minute).

[a] Field entry values taken from the run sheet.

literature, and feasibility to implement in software (see Table 1). To perform metric calculations, sliding windows of 200 RRIs were employed for short-term variability analysis, and 800 RRIs, for long-term variability analysis. Moreover, specific values of $m=2$, $r=6$, and $s=4$ were chosen based upon their performance over repeated animal studies conducted at this Institute. Choice of window sizes was governed by previous work involving heart-rate complexity in trauma patients [16–18], which determined that sample entropy was relatively unaffected by a decrease in number of RRIs (down to a data set of 200 RRIs) whereas multiscale entropy still required a larger set of 800 RRIs for reliable computation.

After obtaining all RRI sequences, metrics were calculated, and mean values and standard deviations were obtained for LSI and non-LSI patient groups. Normality was not assumed for mean values across all records within each group due to the small sample size. Therefore, means between groups were compared using Wilcoxon (non-parametric) statistical tests for true RRI sequences, modified RRI sequences, and detected RRI sequences.

To address how simulated noise could potentially affect the ability of measures to discriminate between LSI and non-LSI trauma patients, noise effects on relationships between mean values across records corresponding to trauma patients who received at least one LSI and those who received none were assessed qualitatively by comparing the selected measures obtained for true RRI sequences, modified RRI sequences, and detected RRI sequences. Similarly, to address how simulated noise could potentially affect the ability of measures to monitor trauma patients in real time, noise effects on HRV and HRC values over time for individual records were assessed qualitatively by analyzing the selected measures as a function of time.

3. Results and discussion

The demographics of the 108 patients included in this study are depicted in Table 2. Quartiles were established for age. Race and age were not statistically different between those patients who received at least one LSI and those who received none. Likewise, increasing patient age did not increase the frequency of an LSI in this sample/study. Of these 108 patients, 26 (24%) did not require an LSI. The other 82 (76%) patients received a total of 142 LSIs. 61% (87) of the LSIs were performed pre-hospital and 39% (55) in the emergency room. Interventions consisted of the following: 73 endotracheal intubations, 30 transfusions, 18 tube thoracostomies, 9 cardiopulmonary resuscitations, 8 needle decompressions, 1 angio-embolization, 2 cricothyrotomies, and 1 thoracotomy. Importantly, the demographics of the chosen population included heart rates ranging from 20 to 156 beats per minute, systolic blood pressures ranging from 80 to 170 mm Hg, diastolic blood pressures ranging from 44 to 110 mm Hg, and various types of injuries and LSIs. This cohort provided the ECG morphology for challenging the RWD algorithm as well as the RRI sequences for noise simulation and analysis.

The performances of the RWD algorithm against all 108 records and against records of individual groups are shown in Table 3. Given the demographics in Table 2, the RWD algorithm achieved an overall sensitivity of 91.8% and positive predictive value of 92.2%. A comparison of sensitivity and positive predictive value between individual groups indicated that performances were not statistically

Table 3
Performance of R-wave detection against ECG records.

# Records	# Beats	TP	FP	FN	Se (%)	+P (%)
108	214,823	197,257	16,603	17,566	91.8	92.2
82 (LSI)	164,786	151,143	12,851	13,643	91.7	92.2
26 (NLSI)	50,037	46,114	3752	3923	92.2	92.5

TP, true positive; FP, false positive; FN, false negative; Se, sensitivity; +P, positive predictive value; LSI, life-saving intervention; NLSI, non-life-saving intervention.

different and that resulting metrics could be compared without bias in RWD performance. In addition, means of selected metrics, standard deviations, and *p*-values obtained via Wilcoxon tests for LSI and non-LSI patient groups are shown in Table 4. Sample entropy (SampEn), quadratic sample entropy (QSE), and multiscale entropy (MSE) values for patients who received at least one LSI were consistent with the fact that this group often has lower HRC than patients who did not receive any LSI [16–18]. Similarly, higher fractal scaling exponent (α), autocorrelation coefficient [$A(k, \tau)$], and degree of non-stationarity (StatAv) values for the LSI patient group were consistent with the fact that this group often has concomitant illnesses [19–23]. Importantly, Table 4 demonstrated that given RWD performances in Table 3, SampEn, QSE, MSE, and $A(k, \tau)$ could still be used to discriminate between different populations. A further implication was that calculation of these metrics could be useful for risk stratification, when simulated noise via detection of an RRI sequence from the ECG (i.e., actual RWD performance) exceeded sensitivities and positive predictive values shown in Table 3.

Table 4
Comparison of detection-derived mean values between patient groups.

Measure	LSI[a] patients (n=82)		NLSI[b] patients (n=26)		*p*-value
	Mean	Std dev	Mean	Std dev	
Sample entropy	0.833	0.303	1.108	0.351	0.002
Quadratic sample entropy	2.989	0.857	3.549	0.736	0.004
Multiscale entropy	0.546	0.427	0.975	0.553	0.001
Poincaré variability ratio	0.508	0.247	0.448	0.158	0.370
Fractal scaling exponent	0.963	0.196	0.897	0.151	0.101
Autocorrelation coefficient	0.112	0.080	0.063	0.046	0.004
Degree of non-stationarity	0.709	0.180	0.634	0.168	0.039
Standard deviation	45.806	60.646	40.399	25.712	0.309
Successive differences	44.520	70.281	35.661	31.291	0.454

[a] LSI denotes life-saving intervention.
[b] NLSI denotes non-LSI.

Means of selected metrics for noise simulated via decimation, concatenation, and division are shown in Figs. 1–3, respectively. In each figure, red diamonds were associated with means of LSI patients, whereas blue squares were associated with means of non-LSI patients. Furthermore, the leftmost points of each plot corresponded to means of true RRI sequences. Decreasing values

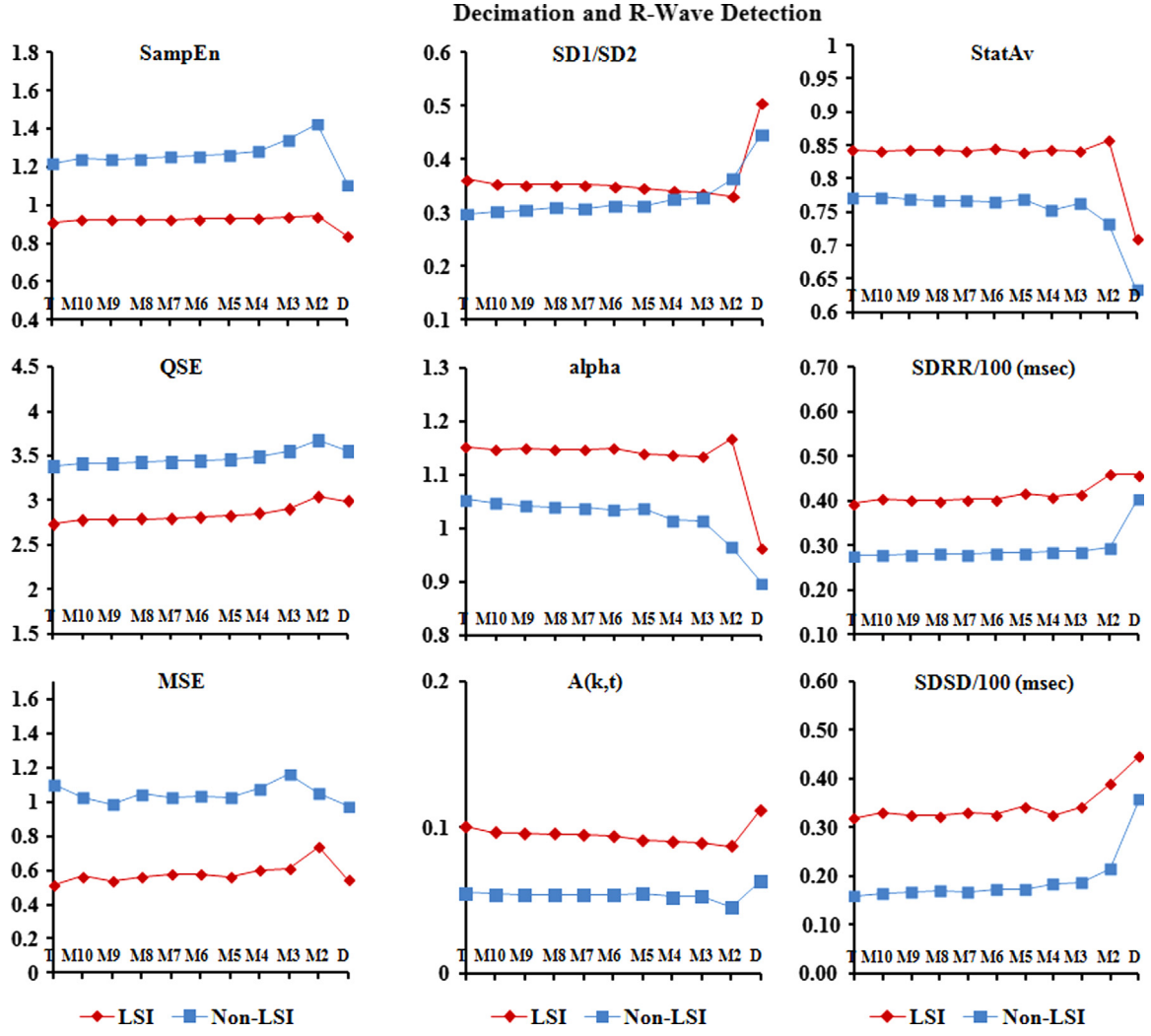

Fig. 1. Means of selected metrics for noise simulated via decimation and R-wave detection. To address how noise could potentially affect the ability of measures to discriminate between LSI and non-LSI trauma patients, noise was simulated by decimating true RRI sequences at every *i*th multiple (M_i) for i=10, 9, …, 2 and by detecting RRI sequences using the ECG waveform. Afterwards, metrics were calculated for true RRI sequences (T), modified RRI sequences (M_{10}, M_9, M_8, …, M_2), and detected RRI sequences (D). Mean values were then obtained for LSI and non-LSI patient groups. Decreasing values of M denoted increasing percentages of a true RRI sequence affected by noise. Decimation was chosen to partially randomize a true sequence in a controlled manner, thereby ensuring that newly formed sequences could be related to the original sequence. It was also used to simulate missing or dropped data. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

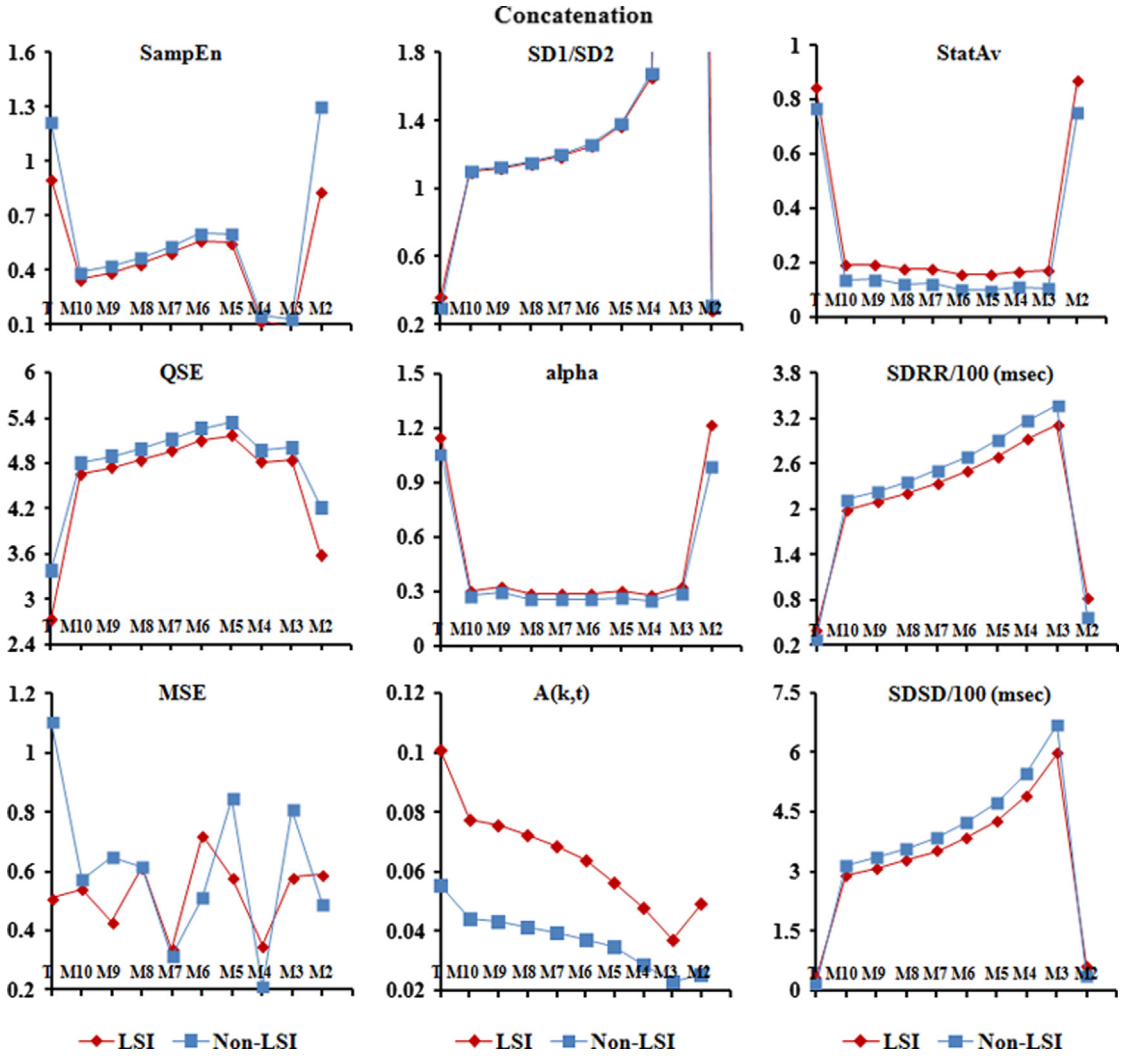

Fig. 2. Means of selected metrics for noise simulated via concatenation. To address how noise could potentially affect the ability of measures to discriminate between LSI and non-LSI trauma patients, noise was simulated by concatenating adjacent RRIs of true RRI sequences at every *i*th multiple (M_i) for $i=10, 9, \ldots, 2$. Afterwards, metrics were calculated for true RRI sequences (T) and modified RRI sequences ($M_{10}, M_9, M_8, \ldots, M_2$). Mean values were then obtained for LSI and non-LSI patient groups. Decreasing values of M denoted increasing percentages of a true RRI sequence affected by noise. Concatenation of selected RRIs was used to simulate missed beats and arrhythmias. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

of M denoted increasing percentages of a true RRI sequence affected by noise.

Although sequences corresponding to M_2 and M_3 in Fig. 1 produced means that deviated most from the true means, in general, increasing simulated noise via decimation did not affect relationships between means of patients groups. An exception was for the Poincaré variability ratio (SD1/SD2) in Fig. 1, where mean values were reversed at M_2.The results in Fig. 1 suggested that for a "lossy" data channel represented by simulated noise via decimation, if the lossy rate was no more than 25% (i.e., between M_9 and M_4), selected metrics other than SD1/SD2 in Table 1 could still be reliable for discriminating between different trauma patient groups.

In Figs. 2 and 3, almost all means of modified RRI sequences differed from true means by over 20%, and in some cases (e.g., SampEn, QSE, and MSE), distances between means of patient groups differed from true distances by over 50%. For SD1/SD2 and α in Fig. 2, there was not a clear distinction between means of patient groups corresponding to modified sequences. Moreover, for the standard deviation of RRIs (SDRR) and standard deviation of successive RRI differences (SDSD) in Figs. 2 and 3, mean values were reversed at all M_i for $i=10, 9, \ldots, 2$. For SampEn under simulated noise via concatenation in Fig. 2, relationships between mean values were preserved. However, neither SampEn in Fig. 3 nor MSE in Figs. 2 and 3 was able to preserve relationships completely. (Consider MSE in Fig. 3 at M_{10}.) Notably, in Figs. 2 and 3, QSE, $A(k, \tau)$, and StatAv were able to reliably discriminate between LSI and non-LSI patient groups.

Considering all noise models, then, results demonstrated the overall effectiveness of QSE and $A(k, \tau)$ to stratify patients and discriminate between different populations. In addition, results indicated that an increasing presence of false positives (motion artifacts, ectopic beats, and arrhythmias) and false negatives (missed beats, arrhythmias) could not only distort the true means of HRV and HRC metrics in Table 1 but also mask the ability of metrics to discriminate reliably. Most vulnerable to these noise effects were SDRR, SDSD, MSE, and SampEn.

Selected metrics as a function of time for noise simulated via decimation, concatenation, and division are shown in Figs. 4–6, respectively. The plotted sequences belonged to one patient record. Longer sequences were associated with true RRI sequences, while shorter sequences were associated with modified RRI sequences. Decreasing values of M denoted increasing percentages of a true RRI sequence affected by noise as well as decreasing lengths of sequences. From a visual perspective, increasing

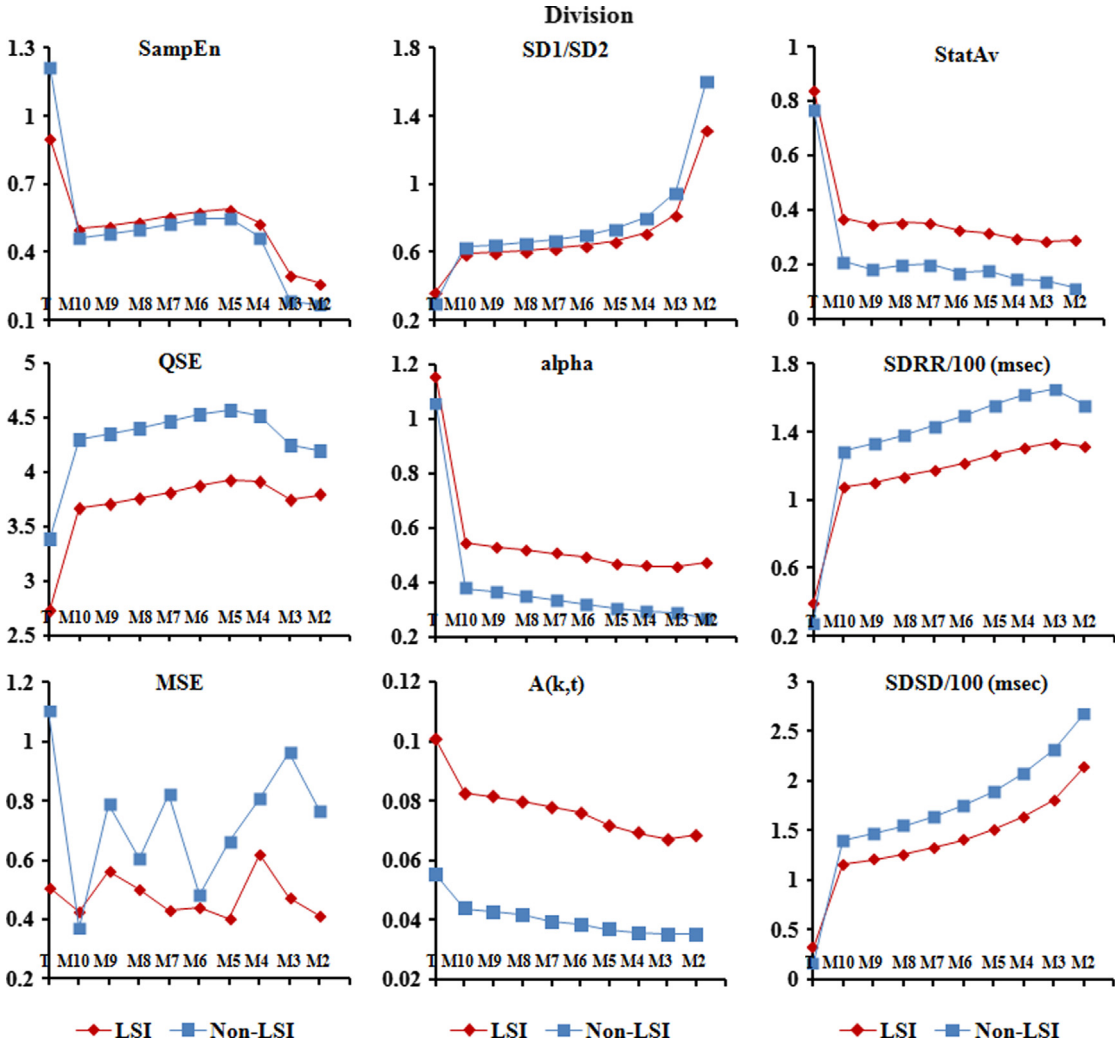

Fig. 3. Means of selected metrics for noise simulated via division. To address how noise could potentially affect the ability of measures to discriminate between LSI and non-LSI trauma patients, noise was simulated by dividing RRIs of true RRI sequences at every ith multiple (M_i) for $i=10, 9, \ldots, 2$. Division of RRIs by two was only performed when the resulting RRIs were each greater than 250 ms in order to ensure that newly created beats did not fall within a refractory period of 220 ms. Afterwards, metrics were calculated for true RRI sequences (T) and modified RRI sequences ($M_{10}, M_9, M_8, \ldots, M_2$). Mean values were then obtained for LSI and non-LSI patient groups. Decreasing values of M denoted increasing percentages of a true RRI sequence affected by noise. Division of selected RRIs was used to simulate ectopic beats, motion artifacts, and arrhythmias. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

simulated noise via decimation did not affect the shapes, trends, and magnitudes of measures as much as increasing simulated noise via concatenation and division for this record. The only exception was for $A(k, \tau)$ in Figs. 4–6, where plots remained similar in shape and magnitude. More apparent in Fig. 4 was the fact that the increasing noise shifted certain features (peaks, valleys) forward in time in Fig. 4. Based on results in Fig. 4, for a "lossy" data channel represented by simulated noise via decimation, selected metrics could still be reliable for monitoring trauma patients in real time. If the magnitudes of SampEn, QSE, and MSE in Figs. 5 and 6 could be scaled appropriately to compensate for noise, perhaps, these measures could also be used for reliable monitoring of the trauma patient. Given the noise models in this study, $A(k, \tau)$ proved to be a viable metric for monitoring trauma patients over time.

4. Conclusion

This study was the first to investigate the impact of noise on the reliability of HRV and HRC metrics to discriminate between different trauma patients and monitor individual trauma patients over time. Using four different noise models and nine different measures, HRV and HRC means were compared between patients who received LSIs and patients who received none. In addition, for one record, HRV and HRC values were compared and analyzed as a function of time.

For the patient cohort in this study, all entropy and autocorrelation measures had power to discriminate between different populations despite suboptimal RWD performance (approximately 92%). A further implication was that calculation of these metrics could be useful for risk stratification and therefore, permit early recognition of needs for LSIs in patients, when RWD performances exceeded such sensitivities and positive predictive values. Additional research will be required to investigate when detection performances could hamper the ability of HRV and HRC metrics to perform risk stratification and how varying detection performances could affect the ability of metrics to monitor patients over time.

Given the simulated noise models in this study, QSE and $A(k, \tau)$ were overall reliable for discriminating between different populations. SDRR, SDSD, MSE, and SampEn failed to discriminate properly in some

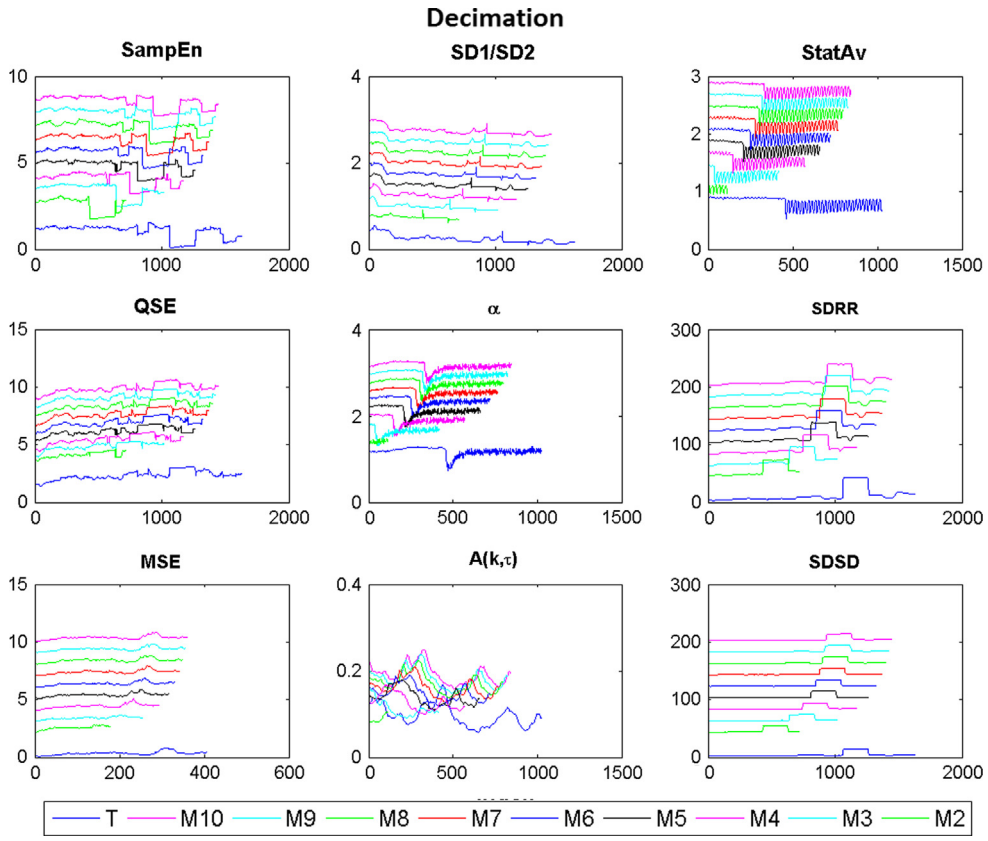

Fig. 4. Selected metrics as a function of time for noise simulated via decimation. To address how simulated noise could potentially affect the ability of measures to monitor trauma patients in real time, noise was simulated by decimating a true RRI sequence corresponding to one patient record at every ith multiple (M_i) for $i=10, 9, \ldots, 2$. Afterwards, metrics were calculated for the true RRI sequence (T) and modified RRI sequences ($M_{10}, M_9, M_8, \ldots, M_2$) and plotted as functions of time (RRIs in ms).

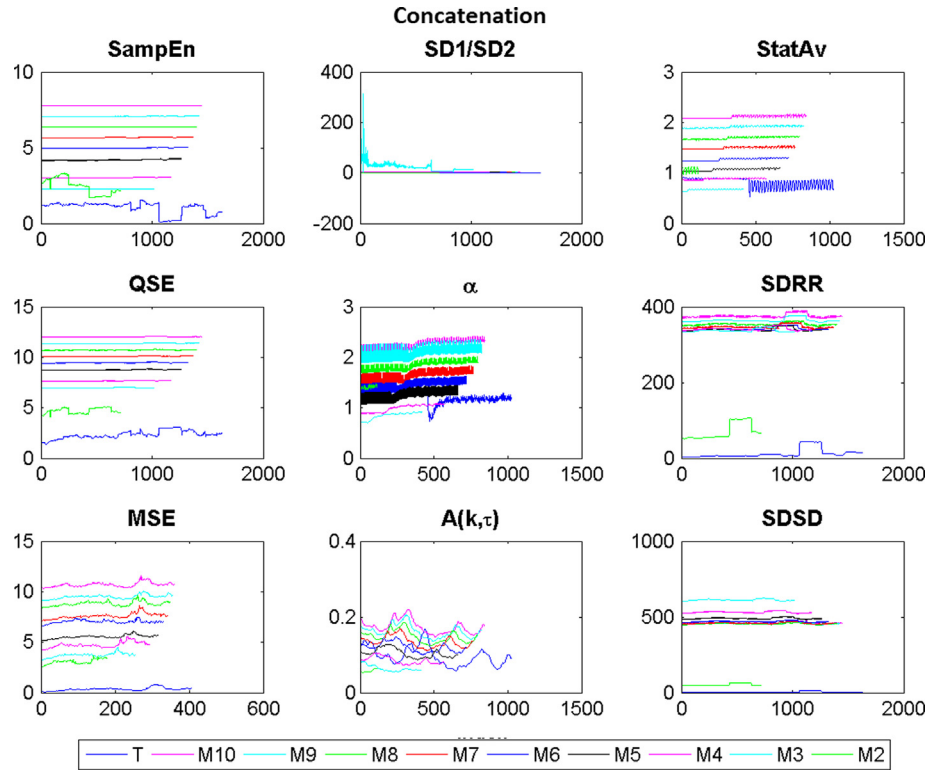

Fig. 5. Selected metrics as a function of time for noise simulated via concatenation. To address how simulated noise could potentially affect the ability of measures to monitor trauma patients in real time, noise was simulated by concatenating adjacent RRIs of a true RRI sequence corresponding to one patient record at every ith multiple (M_i) for $i=10, 9, \ldots, 2$. Afterwards, metrics were calculated for the true RRI sequence (T) and modified RRI sequences ($M_{10}, M_9, M_8, \ldots, M_2$) and plotted as functions of time (RRIs in ms).

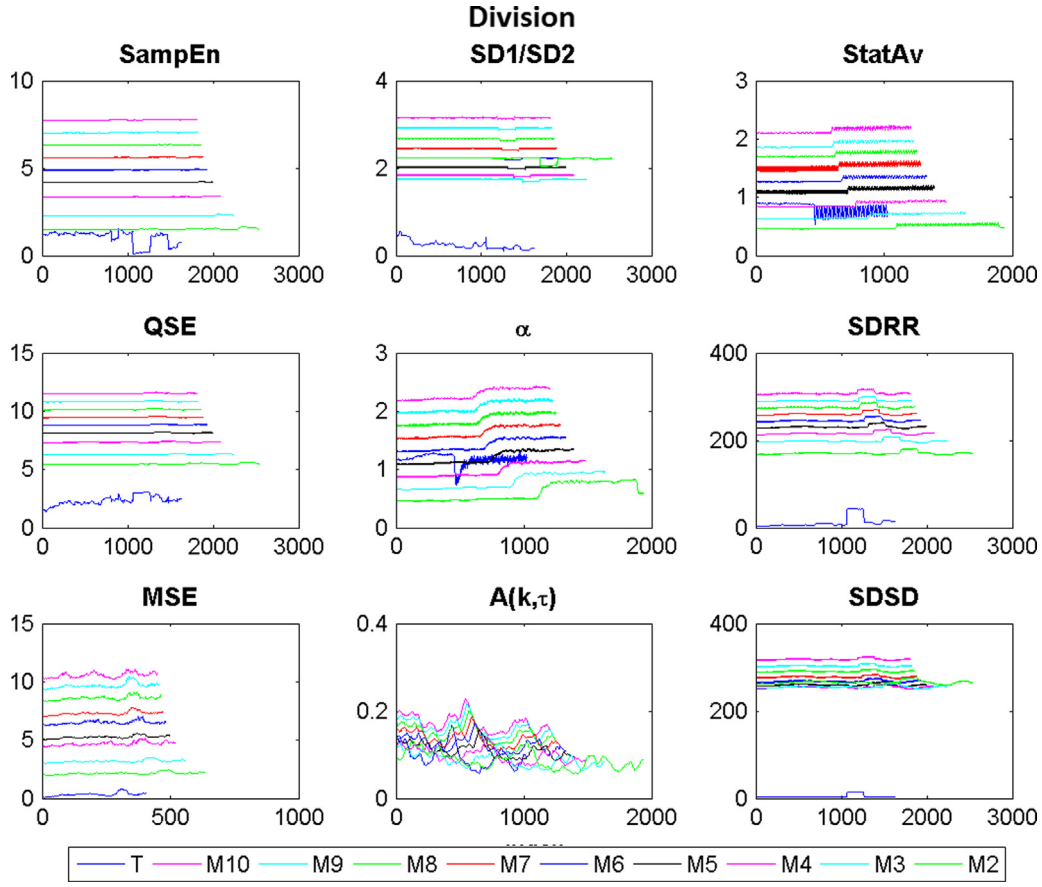

Fig. 6. Selected metrics as a function of time for noise simulated via division. To address how simulated noise could potentially affect the ability of measures to monitor trauma patients in real time, noise was simulated by dividing RRIs of a true RRI sequence corresponding to one patient record at every ith multiple (M_i) for $i=10, 9, \ldots, 2$. Division of RRIs by two was only performed when the resulting RRIs were each greater than 250 ms in order to ensure that newly created beats did not fall within a refractory period of 220 ms. Afterwards, metrics were calculated for the true RRI sequence (T) and modified RRI sequences (M_{10}, M_9, M_8, …, M_2) and plotted as functions of time (RRIs in ms).

situations. An increasing presence of false positives (motion artifacts, ectopic beats, and arrhythmias) and false negatives (missed beats, arrhythmias) demonstrated that HRV and HRC metrics could be distorted as well as lose the ability to discriminate between patient groups reliably. Most vulnerable to these noise effects were SDRR, SDSD, MSE, and SampEn.

For one record in this study, $A(k, \tau)$ was reliable for monitoring due to its ability to maintain its shape, magnitude, and characteristics under the simulated noise models. If proper scaling was applied to the entropy measures in order to compensate for noise effects, it is possible that these metrics could be reliable for monitoring also. Using a similar approach, additional research will be needed to understand a general behavior of metrics for monitoring a diverse set of patients in the presence of noise.

Although this study showed that the calculation of certain HRV and HRC metrics may be less susceptible to simulated noise, and therefore, may be integrated into a real-time software program for decision support and triage in trauma patients, it had one major limitation. Because the dataset contained a small sample size of patients from a specific population, results could not be generalized for healthy versus unhealthy patients. Additional research using a larger patient cohort that includes various diagnoses and physiologic states will be required in order to gain a general understanding of the impact of noise on the reliability of HRV and HRC metrics for use in the clinical environment. A strategy similar to this study could be applied. Development of a methodology and explanations of noise effects will continue to be topics for future research.

Conflict of interest statement

Drs. Batchinsky, Cancio, and Salinas have intellectual property (patent pending, no royalties) dealing with the use of heart-rate complexity in combination with other data to identify critical illness in patients. A patent application will be filed by Mr. Liu with the U.S. Patent Office for the R-wave detection software algorithm.

Acknowledgments

This work was supported by the U.S. Army Combat Casualty Care Research Program. The authors thank Corina Necsoiu and Kerfoot Walker III, who performed the manual verification of all R waves for 108 trauma patient records in the Trauma Vitals database.

Disclaimer: The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the Department of the Army or the Department of Defense.

Appendix A

A.1. Sample, quadratic sample, multiscale entropy

The sample entropy, SampEn(m, r), equals the negative natural logarithm of the conditional probability that two epochs similar for m RRIs remain similar at the next RRI, given a sequence of N

RRIs and excluding self-matches [24]. Here, similarity means that RRIs differ by no more than some tolerance r, be it seconds or milliseconds.

For clarity, sample entropy was computed by the following equations:

$$\mathrm{SampEn}(m,r,N) = -\ln(A/B), \tag{1}$$

$$B = [(N-m-1)/2]\sum_{i=1}^{N-m} B_i^r(m), \tag{2}$$

$$A = [(N-m-1)/2]\sum_{i=1}^{N-m} A_i^r(m). \tag{3}$$

In other words, for a sequence of N intervals, if $x_m(i)$ is an epoch of m consecutive intervals starting at index i and running from $i=1, \ldots, N-m$, then $B_i^r(m)$ denotes the number of epochs $x_m(j)$ within r of $x_m(i)$, for $i \neq j$, multiplied by $(N-m-1)^{-1}$, and $A_i^r(m)$ denotes the number of epochs $x_{m+1}(j)$ within r of $x_{m+1}(i)$, for $i \neq j$, multiplied by $(N-m-1)^{-1}$ [24–26].

The quadratic sample entropy, QSE(m, r), equals SampEn(m, r) normalized by $2r$. Obtaining a coarse-grained RRI sequence from the original sequence and calculating SampEn(m, r) produces the multiscale entropy, MSE(m, r, s), where s denotes a scaling factor used in the coarse-graining equation:

$$y_j = \frac{1}{s}\sum_{i=(j-1)s+1}^{js} x_i \ , \tag{4}$$

$\forall$ integers j such that $1 \le j \le N/s$, x_i denotes the ith RRI of the original sequence, and y_j denotes the jth RRI of the coarse-grained sequence [25,26].

A.2. Poincaré variability ratio

The Poincaré plot takes an RRI sequence and plots each RRI against the following RRI. A quantitative analysis of the plot's geometry involves the calculation of two standard deviations, SD1 and SD2, respectively, describing the short-term and long-term RRI variabilities of a sequence, as well as calculation of the variability ratio SD1/SD2. Because SD1 and SD2 correspond to the minor and major axes, respectively, of an ellipse fitted over the scatter plot, they can be determined numerically by calculating the standard deviations of distances from each point on the plot to the lines $\mathrm{RRI}_{i+1}=\mathrm{RRI}_i$ and $\mathrm{RRI}_{i+1}=-\mathrm{RRI}_i+2\mathrm{RRI}_{mean}$, respectively. Alternatively, they can be obtained from the following equations:

$$\mathrm{SD1}^2 = \mathrm{var}\left(\frac{1}{\sqrt{2}}\mathrm{RRI}_i - \frac{1}{\sqrt{2}}\mathrm{RRI}_{i+1}\right) = \frac{1}{2}\mathrm{SDSD}^2, \tag{5}$$

and

$$\mathrm{SD2}^2 = 2\mathrm{SDRR}^2 - \frac{1}{2}\mathrm{SDSD}^2, \tag{6}$$

$\forall$ $i=1, \ldots, N$, where var($\cdot$) denotes the variance function, SDRR denotes the standard deviation of RRIs, SDSD denotes the standard deviation of successive RRI differences, and N denotes the number of RRIs in the sequence window [27].

A.3. Fractal scaling exponent

Detrended fluctuation analysis (DFA) quantifies the long-range correlation of RRI sequences by estimating the fractal scaling exponent α within a sequence [19–21]. The following steps are used to perform DFA on an RRI sequence with N RRIs:

1. Integrate the RRI sequence to obtain $y(k)$, where k denotes the index of an RRI in the original sequence.
2. Divide $y(k)$ into boxes of length n, for some box size n.
3. For each box, calculate a least squares line $y_n(k)$ best fitted to $y(k)$, the line itself representing the trend in that box.
4. Calculate the root-mean-square (RMS) fluctuation of the integrated and detrended sequence using the equation

$$F(n) = \sqrt{\frac{1}{N}\sum_{k=1}^{N}[y(k)-y_n(k)]^2} \ . \tag{7}$$

5. Repeat steps 2–4 over all time scales (box sizes) in order to characterize the relationship between log $F(n)$ and log n.
6. Obtain the slope α of the line relating log $F(n)$ and log n.

Here, the linear relationship between $F(n)$ and n on a log–log scale describes the degree of power law (fractal) scaling within $y(k)$. In other words, the scaling exponent α indicates the roughness of the original RRI sequence. The larger the value of α, the smoother the RRI sequence. According to this method, $\alpha \approx 0.5$ implies that a sequence resembles white Gaussian noise (a totally random sequence), while $\alpha \approx 1.0$ implies that the sequence resembles a fractal-like time series. Analogously, $\alpha \approx 1.5$ implies that a sequence resembles Brownian noise with decreasing power in its high frequency components [19–21].

In addition, suitable values of n for estimating the scaling exponent α include $4 \le n \le N/4$ uniformly distributed over log n. Another form of analysis is to estimate a short-term scaling exponent α_1 for $4 \le n \le 16$ and a long-term scaling exponent α_2 for $16 \le n \le 64$. Although these techniques make DFA more applicable to both large and relatively short data sets, in general, larger data lengths and the editing of ectopic beats will improve data analysis [19–21].

A.4. Autocorrelation coefficient

The autocorrelation coefficient utilizes the histogram. Given a defined RRI window, this coefficient calculates the overlap of the histograms for two consecutive RRI windows according to the equation

$$A(k,\tau) = \sum_{bin=1}^{N_{bins}} [p_{bin}(k)\ \ p_{bin}(k+\tau)] \ , \tag{8}$$

where k denotes the starting index of an RRI in the original windowed sequence, τ denotes a shift from the starting index, $p_{bin}(k)$ denotes the probability that an RRI falls into the designated bin of the histogram probability distribution, and N_{bins} denotes the number of bins in the histogram. Since the coefficient $A(k, \tau)$, sometimes referred to as the similarity of distributions (SOD), measures a probability, it falls into the range of [0, 1]. In particular, it tends toward 0 if the RRI histograms have non-zero values dispersed across all bins, and it tends towards 1 if the histograms have similar peaks. Hence, the SOD measures the predictability and stability of the heart's dynamics at a specific time [22].

A.5. Degree of non-stationarity

The index StatAv measures the degree of non-stationarity within an RRI sequence. Formally, StatAv breaks up a given RRI window into 40 epochs, computes the sample mean of each epoch, and then calculates the ratio between the standard deviation of the 40 means and the SDRR. Given N RRIs in an RRI window, this can be expressed mathematically as follows:

$$\mathrm{StatAv} = \mathrm{SDAV}/\mathrm{SDRR} \ , \tag{9}$$

where SDAV denotes the standard deviation of $M_1, M_2, \ldots, M_{40}$ and M_k denotes the sample mean of an epoch formed inclusively from RRI_i between $i=N\cdot(k-1)/40+1$ and $N\cdot k/40$, $\forall$ $k=1, \ldots, 40$. Based upon

this definition, smaller StatAv values imply more stationarity. Thus, StatAv quantifies the tendency of the mean to vary with time [23].

References

[1] A. Voss, S. Schulz, R. Schroeder, M. Baumert, P. Caminal, Methods derived from nonlinear dynamics for analysing heart rate variability, Philos. Trans. R. Soc. A 367 (2009) 277–296.

[2] U.R. Acharya, K.P. Joseph, N. Kannathal, C.M. Lim, J.S. Suri, Heart rate variability: a review, Med. Biol. Eng. Comput. 44 (2006) 1031–1051.

[3] Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Heart rate variability: standards of measurement, physiological interpretation and clinical use, Circulation 93 (1996) 1043–1065.

[4] M. Malik, A.J. Camm, Heart rate variability: from facts to fancies, J. Am. Coll. Cardiol. 22 (1993) 566–568.

[5] Q. Li, R.G. Mark, G.D. Clifford, Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter, Physiol. Meas. 29 (2008) 15–32.

[6] P.S. Hamilton, W.J. Tompkins, Quantitative investigation of QRS detection rules using the MIT/BIH Arrhythmia Database, IEEE Trans. Biomed. Eng. 33 (1986) 1157–1165.

[7] G.M. Friesen, T.C. Jannett, M.A. Jadallah, S.L. Yates, S.R. Quint, H.T. Nagle, A comparison of the noise sensitivity of nine QRS detection algorithms, IEEE Trans. Biomed. Eng. 37 (1990) 85–98.

[8] I.I. Christov, Real time electrocardiogram QRS detection using combined adaptive threshold, Biomed. Eng. Online 3 (2004) 1–9.

[9] M.L. Ryan, C.M. Thorson, C.A. Otero, T. Vu, K.G. Proctor, Clinical applications of heart rate variability in the triage and assessment of traumatically injured patients, Anesthesiol. Res. Pract. 2011 (2011) 1–8.

[10] J.R. Moorman, W.A. Carlo, J. Kattwinkel, R.L. Schelonka, P.J. Porcelli, C.T. Navarrete, E. Bancalari, J.L. Aschner, M. Whit Walker, J.A. Perez, C. Palmer, G.J. Stukenborg, D.E. Lake, T. Michael O'Shea, Mortality reduction by heart rate characteristic monitoring in very low birth weight neonates: a randomized trial, J. Pediatr. 6 (2011) 900–906.

[11] J.R. Moorman, J.B. Delos, A.A. Flower, H. Cao, B.P. Kovatchev, J.S. Richman, D.E. Lake, Cardiovascular oscillations at the bedside: early diagnosis of neonatal sepsis using heart rate characteristics monitoring, Physiol. Meas. 32 (2011) 1821–1831.

[12] G. Green, B. Bradley, A. Bravi, A.J. Seely, Continuous multiorgan variability analysis to track severity of organ failure in critically ill patients, J. Crit. Care 28 (2013) 879.

[13] A. Bravi, G. Green, A. Longtin, A.J. Seely, Monitoring and identification of sepsis development through a composite measure of heart rate variability, PLoS One 7 (2012) e45666.

[14] A. Bravi, A. Longtin, A.J. Seely, Review and classification of variability analysis techniques for clinical applications, Biomed. Eng. Online 10 (2011) 1–27.

[15] N.T. Liu, A.I. Batchinsky, L.C. Cancio, W.L. Baker, J. Salinas, Development and validation of a novel fusion algorithm for continuous, accurate, and automated R-wave detection and calculation of signal-derived metrics, J. Crit. Care 28 (2013) 885.

[16] A.I. Batchinsky, L.C. Cancio, J. Salinas, T. Kuusela, W.H. Cooke, J.J. Wang, M. Boehme, V.A. Convertino, J.B. Holcomb, Prehospital loss of R-to-R interval complexity is associated with mortality in trauma patients, J. Trauma 63 (2007) 512–518.

[17] L.C. Cancio, A.I. Batchinsky, J. Salinas, T. Kuusela, V.A. Convertino, C.E. Wade, J.B. Holcomb, Heart-rate complexity for prediction of prehospital lifesaving interventions in trauma patients, J. Trauma 65 (2008) 813–819.

[18] A.I. Batchinsky, J. Salinas, T. Kuusela, C. Nescoiu, J. Jones, L.C. Cancio, Rapid prediction of trauma-patient survival by analysis of heart-rate complexity: impact of reducing dataset size, Shock 32 (2009) 565–571.

[19] A.L. Goldberger, B.J. West, Application of nonlinear dynamics to clinical cardiology, Ann. N. Y. Acad. Sci 504 (1987) 195–213.

[20] A.L. Goldberger, L.A. Amaral, J.M. Hausdorff, P.C. Ivanov, C.K. Peng, H.E. Stanley, Fractal dynamics in physiology: alterations with disease and aging, Proc. Natl. Acad. Sci. U.S.A. 99 (Suppl. 1) (2002) 2466–2472.

[21] K.K. Ho, G.B. Moody, C.K. Peng, J.E. Mietus, M.G. Larson, D. Levy, A.L. Goldberger, Predicting survival in heart failure case and control subjects by use of fully automated methods for deriving nonlinear and conventional indices of heart rate dynamics, Circulation 96 (1997) 842–848.

[22] M. Zochowski, K. Winkowska-Nowak, A. Nowak, G. Karpinski, A. Budaj, Autocorrelations of R–R distributions as a measure of heart variability, Phys. Rev. E 56 (1997) 3725–3727.

[23] S.M. Pincus, R.R. Viscarello, Approximate entropy: a regularity measure for fetal heart rate analysis, Obstet. Gynecol. 79 (1992) 249–255.

[24] J.S. Richman, J.R. Moorman, Physiological time series analysis using approximate entropy and sample entropy, Am. J. Physiol. Heart Circ. Physiol. 278 (2000) H2039–H2049.

[25] M. Costa, A.L. Goldberger, C.K. Peng, Multiscale entropy analysis of biological signals, Phys. Rev. E 71 (2005) 021906.

[26] M. Costa, A.L. Goldberger, C.K. Peng, Multiscale entropy analysis of physiologic time series, Phys. Rev. Lett. 89 (2002) 062102.

[27] M. Brennan, M. Palaniswami, P. Kamen, Do existing measures of Poincarè plot geometry reflect nonlinear features of heart rate variability? IEEE Trans. Biomed. Eng. 48 (2001) 1342–1347.